

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-167901

(43)Date of publication of application : 13.06.2003

(51)Int.Cl.

G06F 17/30

G06F 17/60

G10K 15/02

(21)Application number : 2001-365215

(71)Applicant : KDDI CORP

(22)Date of filing : 29.11.2001

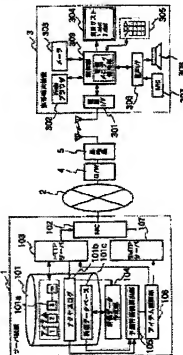
(72)Inventor : HOASHI KEIICHIRO
MURAMATSU SHIGEKI
TANAKA TAKUYA
INOUE NAOKI

(54) METHOD OF COOPERATIVE FILTERING, DEVICE FOR COOPERATIVE FILTERING AND PROGRAM FOR COOPERATIVE FILTERING

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method of cooperative filtering or the like capable of improving the accuracy of forecast evaluation values while restraining the burden on users.

SOLUTION: An evaluated data preparation part 104 prepares the evaluation data for each of items concerning each of the users according to an access log 101b, and stores them in an evaluation database 101c. A forecast evaluation value calculation part 105 calculates the forecast evaluation values concerning the items that are not evaluated by specific users, according to the evaluation data in the evaluation database 101c. An item recommendation part 106 recommends the items that have high forecast evaluation values as a result of calculation, to the specific users.



(19) 日本特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-167901

(P2003-167901A)

(43) 公開日 平成15年6月13日(2003.6.13)

(51) Int. Cl. ⁷	識別記号	F I	7-23-17 (参考)
G 0 6 F 17/30	3 4 0	G 0 6 F 17/30	3 4 0 A 5 B 0 7 5
17/00	1 5 0	17/00	1 5 0
G 1 0 K 15/02		G 1 0 K 15/02	

特許請求 未請求 請求項の数 5 ○ L (全 10 頁)

(21) 出願番号 特願2001-385215(P2001-385215)
 (22) 出願日 平成13年11月29日(2001.11.29)

(71) 出願人 000208891
 KDD I 株式会社
 東京都新宿区西新宿二丁目3番2号
 (72) 発明者 帆足 吾一郎
 埼玉県上福岡市大原2丁目1番15号 株式会社
 ケイディーディーアイ研究所内
 (72) 発明者 村松 茂樹
 埼玉県上福岡市大原2丁目1番15号 株式会社
 ケイディーディーアイ研究所内
 (74) 代理人 10005-8870
 弁理士 田中 富樹 (外2名)

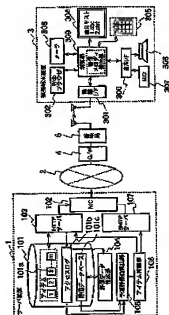
最終頁に続く

(54) 【発明の名称】 協調フィルタリング方法、協調フィルタリング装置及び協調フィルタリングプログラム

(57) 【要約】

【課題】 ユーザの負担を抑えつつ予測評価値の精度を向上させることができる協調フィルタリング方法等を提供する。

【解決手段】 評価データ作成部104は、アクセスログ101bに応じて各々のユーザについて各々のアイテムに対する評価データを作成し、評価データベース101cに格納する。予測評価値算出部105は、評価データベース101c中の評価データに応じて特定のユーザが未評価であるアイテムについての予測評価値を求める。アイテム総覧部106は、求められた予測評価値が高いアイテムを特定のユーザ宛てに総覧する。



特開2003-167901

1

【特許請求の範囲】

【請求項1】 特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリング方法であって、

ユーザの行動履歴に応じた評価対象についての評価を示す暗示的評価データを取得するステップと、
前記暗示的評価データと、前記ユーザの行動による評価の時から経過時間に応じた評価情報を求めるステップと、

前記特定のユーザと前記評価情報の傾向が類似するユーザを選択するステップと、

該選択したユーザが評価している評価対象についての前記評価情報に応じて、前記特定のユーザの当該評価対象についての予測評価値を求めるステップとを有することを特徴とする協調フィルタリング方法、

【請求項2】 前記予測評価値に応じて当該予測評価値に対応する前記評価対象を前記特定のユーザに推薦するステップとを有することを特徴とする請求項1記載の協調フィルタリング方法、

【請求項3】 特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリング装置であって、

評価対象をネットワーク経由でユーザに提供する提供手段と、

前記各ユーザ毎に対する評価対象の提供履歴に応じて各ユーザの評価対象に対する評価を示す暗示的評価データを記録する暗示的評価データ記録手段と、

前記記録した暗示的評価データと、前記ユーザに対する評価対象の提供の時から経過時間に応じた評価情報を求める評価情報算出手段と、

前記特定のユーザと前記評価情報の傾向が類似するユーザを選択するユーザ選択手段と、

該選択したユーザが評価している評価対象についての前記評価情報に応じて、前記特定のユーザの当該評価対象についての予測評価値を求める予測評価値算出手段とを備えることを特徴とする協調フィルタリング装置、

【請求項4】 前記予測評価値算出手段が求めた予測評価値に応じて当該予測評価値に対応する前記評価対象を前記特定のユーザに推薦する推薦手段を備えることを特徴とする請求項3記載の協調フィルタリング装置、

【請求項5】 コンピュータに、特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリングプログラムであって、

ユーザの行動履歴に応じた評価対象についての評価を示す暗示的評価データを取得するステップと、
前記暗示的評価データと、前記ユーザの行動による評価の時から経過時間に応じた評価情報を求めるステップと、

前記特定のユーザと前記評価情報の傾向が類似するユーザを選択するステップと、

(2)

2

該選択したユーザが評価している評価対象についての前記評価情報に応じて、前記特定のユーザの当該評価対象についての予測評価値を求めるステップとを有することを特徴とする協調フィルタリングプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ユーザの興味に合う情報を選択する協調フィルタリング方法、協調フィルタリング装置及び協調フィルタリングプログラムに関する。

【0002】

【従来の技術】膨大な情報の中からユーザの興味に合致する情報を選択する技術として、協調フィルタリングが知られている。

【0003】この協調フィルタリングでは、特定の利用者が未評価である対象について評価値を予測する際に、評価済の対象についての評価の傾向が似ている（すなわち評価値に相関がある）他のユーザの評価値に基いて予測評価値を求めている。

【0004】従来の協調フィルタリングでは、評価値を予測するための基準となる個々のユーザの評価値データとして、主に明示的（explicit）な評価データと、暗示的（implicit）な評価データが用いられている。

【0005】明示的な評価データとは、予め実施したアイテム（コンテンツ等の情報）に対するアンケート等に対して、ユーザが意図的に評価を与えたデータである。この明示的な評価データの一例としてEachMovieが挙げられる。このEachMovieに属する評価データは、図8に示すように、個々のユーザが見た個々の映画（アイテム）を5段階で評価したデータの集合である。

【0006】これに対し、暗示的な評価データとは、ユーザが意識することなく得られた評価データである。この暗示的な評価データは、Webページ等のアクセス履歴、テレビ番組の視聴率調査等によって得られるデータである。このような暗示的な評価データは、図9に示すように、あるアイテムをユーザが参照したか否かを数値（参照した場合には“1”、参照していない場合には“0”）で示したものである。

【0007】協調フィルタリングでは、上述のように得られた評価データに基いて、評価データの傾向が類似しているユーザをグループ化し、グループ内の他のユーザは利用したか当該ユーザが未だ利用していない情報を予測評価値として求め、予測評価値の高いアイテムをユーザに推薦するようになっている。

【0008】

【発明が解決しようとする課題】しかしながら、上述の明示的な評価データに応じた協調フィルタリングでは、得られる予測評価値の精度は高いものの、ユーザが各々のアイテムに対して評価を行うことが前提となっている

(3)

特開2003-167901

3

ため、実際に適用する環境ではユーザの負担が重く、現実的ではない。

【0009】一方、暗示的な評価データを用いる場合には、ユーザに負担をかけるとなく評価データを得ることができ、例えばユーザのアクセス履歴に基づく場合等では、ユーザに対する情報量が少なく、得られる予測評価値の精度が低くなってしまうことがある。

【0010】本発明は、上述のような課題に鑑みてなされたものであり、ユーザの負担を抑えつつ予測評価値の精度を向上させることができる協調フィルタリング方法、協調フィルタリング装置及び協調フィルタリングプログラムを提供することを目的とする。

【0011】

【課題を解決するための手段】上述の問題を解決するために、本発明に係る協調フィルタリング方法は、特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリング方法であって、ユーザの行動履歴に応じた評価対象についての評価を示す暗示的な評価データを取得するステップと、暗示的な評価データと、ユーザの行動による評価の時から経過時間に応じた評価情報を求めるステップと、特定のユーザと評価情報の傾向が類似するユーザを選択するステップと、選択したユーザが評価している評価対象についての評価情報に応じて、特定のユーザの当該評価対象についての予測評価値を求めるステップとを有することを特徴とする。

【0012】評価対象には、例えばネットワーク経由でユーザの端末装置に提供されるニュース等の文字情報、音信メロディ、音楽等のコンテンツ等が含まれる。

【0013】また、暗示的な評価データとは、ユーザが意図的に評価を行わなくても生成されるデータであり、例えばサーバ装置に対するアクセスログ等が含まれる。

【0014】また、本発明に係る協調フィルタリング装置は、特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリング装置であって、評価対象をネットワーク経由でユーザに提供する提供手段と、各ユーザに対する評価対象の提供履歴に応じてユーザの評価対象に対する評価を示す暗示的な評価データを記録する暗示的な評価データ記録手段と、記録した暗示的な評価データと、ユーザに対する評価対象の提供の時から経過時間に応じた評価情報を求める評価情報算出手段と、特定のユーザと評価情報の傾向が類似するユーザを選択するユーザ選択手段と、選択したユーザが評価している評価対象についての評価情報に応じて、特定のユーザの当該評価対象についての予測評価値を求める予測評価値算出手段とを備えることを特徴とする。

【0015】また、本発明に係る協調フィルタリングプログラムは、コンピュータに、特定のユーザが評価していない評価対象についての予測評価値を求める協調フィルタリングプログラムであって、ユーザの行動履歴に応じた評価対象についての評価を示す暗示的な評価デー

4

タを取得するステップと、暗示的な評価データと、ユーザの行動による評価の時から経過時間に応じた評価情報を求めるステップと、特定のユーザと評価情報の傾向が類似するユーザを選択するステップと、選択したユーザが評価している評価対象についての評価情報に応じて、特定のユーザの当該評価対象についての予測評価値を求めるステップとを有することを特徴とする。

【0016】

【発明の実施の形態】本発明は、例えばインターネットを介して、ニュース等の文字情報、音信メロディ、音楽等のコンテンツ（アイテム）をインターネット、電子メール等によって配信する推薦システムに適用することができる。

【0017】（構成）本発明の一実施形態に係る推薦システムは、例えば図1に示すように、アイテムの提供等を行うサーバ装置1と、このサーバ装置1とインターネット等のネットワーク2を介して通信を行うことができる携帯端末装置3とを備えている。

【0018】また、この推薦システムは、サーバ装置1と携帯端末装置3との間の通信プロトコルの交換を行うゲートウェイ装置（GW）4と、携帯端末装置3との間で無線通信を行う基地局5とを備えている。

【0019】（サーバ構成）サーバ装置1は、例えばMPU、メモリ、ハードディスクドライブ（HDD）101、ネットワークインタフェース（NIC）102、入出力手段等を有するパーソナルコンピュータ、ワークステーション等の情報処理装置から構成されている。

【0020】このサーバ装置1では、例えばMPUがHDD101に格納されている制御プログラムを実行することにより、HTTP（Hyper Text Transfer Protocol）サーバ103、評価データ作成部104、予測評価値算出部105、アイテム推薦部106及びSMTP（Simple Mail Transfer Protocol）サーバ107等として機能するようにしている。

【0021】HTTPサーバ103は、ネットワーク2経由で接続される携帯端末装置3等に対して、ニュース等の文字情報、音信メロディ、音楽等のアイテムをHTTPプロトコルによって提供する。

【0022】HDD101には、ユーザに提供し得るアイテムを格納するアイテム格納領域101aが設けられている。このアイテム格納領域101aには、ニュース等の文字情報、音信メロディ、音楽等のアイテムが格納される。

【0023】また、HDD101には、HTTPサーバ103によって、ユーザからのアクセスを示すアクセスログ101bが格納される。

【0024】また、HDD101には、評価データ作成部104が上述のアクセスログ101b内のデータから作成したユーザ毎の評価データが評価データベース101cとして格納される。

(4)

特開2003-167901

5

【0025】評価データ作成部104は、アクセスログ101b中のデータから各々のユーザがアクセスしたアイテムの評価データを作成し、評価データベース101cに格納する。

【0026】予測評価値算出部105は、各ユーザ毎に、評価データベース101c中の評価データに基いてユーザに推薦すべきアイテムを選択するための情報となる予測評価値を求め、アイテム推薦部106に供給する。

【0027】アイテム推薦部106は、供給された予測評価値に基いてユーザに推薦すべきアイテムを決定する。このアイテム推薦部106は、例えばSMTPサーバ107によってユーザに送信するメール、HTTPサーバ103によって提供する推薦するアイテムを示す画面を表示するページ記述データ等によって、ユーザにアイテムを推薦する。

【0028】（情報検索・動作履歴）また、携帯端末装置3は、携帯電話、パーソナル・デジタル・アシスタント（PDA）等の携帯端末装置から構成することができる。この携帯端末装置3は、例えばMPU、メモリ、無線1/F301、表示部304、入力部305、音声1/F306等を有している。この携帯端末装置3では、例えばMPUがメモリに格納されている制御プログラムを実行することにより、Webブラウザ302、メール303、制御部309等が機能するようになっている。

【0029】無線1/F301は、例えば既存のWebブラウザ機能等を備えたデジタルセルラー電話システムと同様の手順で、基地局5との間で無線通信を行うようになっている。この無線1/F301は、基地局5から受信した信号を復調し、受信データとして制御部309に供給すると共に、制御部309からの送信データに応じて送信信号を形成して基地局5に送出している。

【0030】この無線1/F301と制御部309との間のデータは、例えば呼の制御等を行う制御データと、通話用の音声データと、データ通信用のデータ等が多量化されたデータである。制御部309は、音声データ等を多量化した無線1/F301に供給する送信データを形成し、無線1/F301からの受信データを分能して音声データ、データ通信用のデータ等に分離する。

【0031】制御部309は、基地局5との間で呼の設定等の制御を行うようになっている。この制御は、既存のデジタルセルラー電話と同様に、所定の制御データの送受信によって行うようになっている。

【0032】また、送信するデータ通信用のデータは、例えばWebブラウザ302、メール303等が生成し、制御部309に供給する。また、受信したデータ通信用のデータは、制御部309が、Webブラウザ302、メール303に振り分け供給する。

【0033】これらのWebブラウザ302、メール303が送受信するデータ通信用のデータのフォーマット

6

は、G/W4との間で設定された所定のフォーマットに従って構成されている。G/W4は、このフォーマットと、ネットワーク2において用いられているデータのフォーマット（例えばTCP/IP）パケットとの間の交換を行う。これにより、例えばWebブラウザ302とHTTPサーバ103の間、メール303とSMTPサーバ107からのメールを受信するPOP（Post Office Protocol）サーバ等との間で通信を行うことができるようになっている。

【0034】制御部309は、例えばWebブラウザ302によってHTTPサーバ103から取得したHTML（Hyper Text Markup Language）形式あるいはXML（eXtensible Markup Language）形式等のページ記述データに応じた画像データを形成する機能を有している。また、制御部309は、メール303によって取得したメールのメッセージ等に応じた画像データを形成する機能を有している。さらに、制御部309は、発信、着信等の当該携帯端末装置3の動作状態に応じた画像データを形成する機能を有している。表示部304に供給する画像データは、当該携帯端末装置3の動作状態、入力部305を介して入力されるユーザからの指示等に応じて制御部309が適宜選択するようになっている。制御部309は、このように選択した画像データに応じて表示部304を駆動する。これにより、表示部304には、動作状態等に応じた画像が表示される。

【0035】ところで、送信する音声データは、MIC307からの音声信号に応じて音声1/F306が生成し、制御部309に供給する。また、受信した音声データは、制御部309から音声1/F306に供給される。音声1/F306は、マイク（MIC）307からの音声信号のアナログ/デジタル変換、符号化等の処理、制御部309から供給された音声データに応じてスピーカ308の駆動等を実行する。

【0036】また、この携帯端末装置3では、基地局5からの着信に応じてメモリ5を発生させることができるようになっている。このために、制御部309は、着信メモリ5を発生するためのデータを保持する着信メモリ5保持部309aを備えている。この着信メモリ5保持部309aには、例えばメモリ5を音調、長さ、強度等で表すデータ（メモリデータ）が格納されている。制御部309は、このメモリデータに応じて音声データを生成する機能を有している。この生成機能は、メモリデータに応じて音声データを生成するソフトウェアとして実装することもできる。同様の機能を有するハードウェアとして実装することもできる。

【0037】このようにメモリデータに応じて生成された音声データは、着信時に制御部309から音声1/F306に供給される。音声1/F306は供給された音声データに応じてスピーカ308を駆動し、メモリデータに応じた着信メモリ5が発生される。

(5)

特開2003-167901

7

【0038】さらに、この情報端末装置3では、上述のHTTPサーバ103がアイテムとして提供するメロディデータをダウンロードして上述の着信メロディ保持部309aに格納することができるようにしている。

【0039】このダウンロードでは、まず、ユーザが入力部305を操作して、例えばHTTPサーバ103が提供するページ記述データ中で定義されているメロディデータのリストから希望のメロディデータを選択する。このようにユーザから特定のメロディデータが指示されると、Webブラウザ302は、選択されたメロディデータを指示データでHTTPサーバ103宛てに送信する。このようなデータを受信すると、HTTPサーバ103は、指示されたメロディデータをアイテム格納領域101aから読み出し、NIC102経由で情報端末装置3宛てに送信する。このように送信されたデータは、ネットワーク2、G/W4、基地局5、無線1/F301、制御部309を経由してWebブラウザ302に供給される。Webブラウザ302は、供給されたメロディデータを着信メロディ保持部309aに格納する。これにより、HTTPサーバ103からダウンロードした、メロディデータに応じた着信メロディを発生し得る状態となる。

【0040】（動作詳細）ところで、情報端末装置3では、情報端末装置3から表示部304に表示可能な文字数が制限されている。このため、HTTPサーバ103が提供するメロディデータの総数が多くなると、上述のようなダウンロードするメロディデータを選択するための操作は、例えばメロディの分野、アーティスト毎等の複数回の選択が必要になる等、複雑になってしまう。

【0041】このため、この推薦システムでは、予測評価部105が、上述の評価データ作成部104によって作成した評価データベース102cに応じて、ユーザの興味のあるメロディデータ（アイテム）を予測し、予測したアイテムをアイテム推薦部106によってユーザに推薦するようにしている。

【0042】この推薦システムでは、このユーザの興味のあるアイテムの予測は、過去のユーザのアクセス履歴等に時間情報考慮して求めた評価データに基づいて行うようになっている。同様のアイテムを利用するユーザは、当該アイテムを利用する時期によって異なる傾向を有することが考えられる。このため、時間情報を考慮して評価データに基づいて推薦するアイテムの予測を行うことにより、ユーザの傾向をより正確に反映した精度の高い予測を行うことができる。

【0043】また、この推薦システムでは、ユーザにアイテムを推薦する際に、推薦するアイテムを容易にアクセスあるいはダウンロードできるように、当該アイテムにアクセスするための情報（例えば当該アイテムのアドレス等を示す情報等）を一括して提供する。このように、推薦するアイテムのアドレスを示す情報を、ユーザに提

8

供することにより、ユーザが推薦されたアイテムを容易にダウンロードし得る状態となる。

【0044】（時間情報を考慮した評価データの作成）アクセスログ101bにはHTTPサーバ103によって、例えば図2に示すように、個々のアクセスを識別するためのID、HTTPサーバ103に対するアクセスの時刻、アクセス元のIPアドレス、アクセスしたユーザを識別するための識別情報（ユーザID）、アクセス対象のアイテム等が対応付けられて記録されている。

【0045】評価データ作成部104は、予めアクセスログ101bを参照して、各々のユーザ毎に、アクセス（あるいはダウンロード）した個々のアイテムと、各々のアイテムにアクセスした時間情報を考慮した評価値を求め、評価データベース101cに格納する。

【0046】この評価値は、個々のアイテムに対するアクセスの有無を示す数値（アクセスがある場合には“1”、アクセスがない場合には“0”）に、時間情報を示す係数1_iを乗じて求める。

【0047】この係数1_iは、ユーザの現在の興味等を正確に評価すべく、現在の時刻を1とし、時間を遡るに従って値が減少する関数、例えば半減期関数、時間の経過に従って減少する1次関数あるいはさらに高次の関数等を用いて求める。

【0048】係数1_iとして半減期関数を用いる場合には、例えば図3に示すように、式次に従って求めることができる。

【0049】

【数1】

$$E_i = 0.5^{t_n - t_i / \tau}$$

ここで、 t_n は現在の時刻、 t_i はユーザがアイテム1（ $i=1, 2, \dots, m-1, m$ ）にアクセス（あるいはダウンロード）した時刻、 τ は係数1_iの値が0.5になる時間（いわゆる半減期）を示している。

【0050】上述のように、アクセスログ101bから求めた評価データベース中の個々のアイテムの評価値は、例えば図4及び図5に示すように、ユーザが個々のアイテムにアクセスした時刻からの経過時間（アクセス時刻と現在の時刻の差）を反映した値となっている。例えば上述の図2中の符号aに対応する図4及び図5中のユーザID1であるユーザ（以下、単にユーザ1という。）のアイテム2の評価値“0.4”は、これより後にアクセスされた図2中の符号bに対応するユーザID1から2であるユーザ（以下、単にユーザ2という。）のアイテム2の評価値“0.5”より小さくなっている。

【0051】また、予測評価部105は、上述のように求められた評価データベース101c中の各評価値を、例えば学習理論で用いられているメモリベースア

特開2003-167901

9

ルゴリズム (Memory-based Algorithm) あるいはメモリベース推論等の計算式に当てはめ、ユーザが評価していないアイテムについての予測評価値を求める。

【0052】具体的には、このような予測評価値では、例えば図6に示すように、まず、あるユーザを特定のユーザとし、特定のユーザが評価しているアイテムと同じアイテムに対する評価値を有するユーザであって当該特定のユーザの評価値との相関が高いユーザを選択する (S1)。

このユーザの選択では、例えば特定のユーザと他のユーザによる個々のアイテムの評価値を比較し、個々のアイテムの評価値毎の相関を求め、例えば相関が高い場合には高い値として評価する。さらに、ユーザ毎に全てのアイテムについての相関を示す値の総和を求める。このようにして求めた総和が大きいユーザを、例えば総和が大きい順に所定数選択する。選択するユーザの数は、例えばユーザ数と予測評価値の幅の関係を参照で求め、必要な精度、演算負荷等に応じて適宜設定する。

【0053】例えば上述の図5に示す状況では、ユーザ1の評価値の傾向と、ユーザn-2の評価値の傾向が似ている。このため、ユーザn-2が、ユーザ1と相関が高いユーザとして選択される。なお、このユーザn-2の他にも相関が高いユーザが所定数選択される。

【0054】次に、選択した所定数のユーザの内、上述の特定のユーザ (特定ユーザ) が評価していないアイテムについて選択しているユーザを選択し (S2)、当該選択したユーザ (選択ユーザ) による当該アイテムの評価値に応じて特定ユーザの予測評価値を求める (S3)。

【0055】上述のメモリベースアルゴリズムによって求められる予測評価値は、ユーザa (特定ユーザ) のアイテムjに対する予測評価値を $P_{a,j}$ 、とすると、例えば次式で示される。

【0056】

【数2】

$$P_{a,j} = \bar{v}_a + K \sum_{i=1}^n w(a,i) (v_{i,j} - \bar{v}_i)$$

ここで、nは上述のように選択した選択ユーザの数であり、 v_i (バー) はユーザaの評価値の平均 (平均評価値) であり、 $w(a,i)$ はユーザaとユーザi (ある選択ユーザ) の相関を表す値 (相関係数) であり、 $v_{i,j}$ はユーザiのアイテムjに対する評価値である。また、Kは $w(a,i)$ の総和が1となるように、すなわち次式を満たすように設定された正規化係数である。

【0057】

【数3】

$$K \sum_{i=1}^n w(a,i) = 1$$

(6)

10

また、ユーザaとユーザiの相関を表す相関係数 $w(a,i)$ は、例えば次式に示すコサイン値による類似度として求めることができる。

【0058】

【数4】

$$w(a,i) = \frac{\bar{v}_a \cdot \bar{v}_i}{|\bar{v}_a| |\bar{v}_i|}$$

この数4式中、右辺分子のベクトル \bar{v}_a はユーザaの評価値を示すベクトル {ベクトル $v_a = (v_{a,1}, v_{a,2}, \dots, v_{a,n})$ } であり、ベクトル \bar{v}_i はユーザiの評価値を示すベクトル {ベクトル $v_i = (v_{i,1}, v_{i,2}, \dots, v_{i,n})$ } であり、ベクトル v_a ・ベクトル v_i は、ベクトル v_a とベクトル v_i の内積である。また、右辺分母の1ベクトル v_a 、1及び1ベクトル v_i は各々ベクトル v_a 及びベクトル v_i のユークリッド長である。

【0059】このようなメモリベースアルゴリズムによって、例えば上述の図5に示す状況では、ユーザ1と相関が高いユーザとして選択されたユーザn-2等によるアイテムmの評価値から求められたユーザ1の予測評価値が1.0となっている。

【0060】ところで、上述の数2式によって求められる予測評価値は、数2式の右辺第2項に示されるように、特定ユーザとの相関係数に応じて個々の選択ユーザの評価値を重み付けして加算した値に応じて求められているため、相関の高いユーザの評価値が予測評価値により大きく反映されることになり、予測評価値の幅が高くなる。

【0061】なお、上述の数2式の代わりに、単純に選択ユーザのアイテムに対する評価値の平均を予測評価値とすることもできる。この場合、上述の数2式によって予測評価値を求める場合に比較して精度は低下するものの、予測評価値を求めるための演算負荷を低減することができる。

【0062】以上のように特定ユーザの未評価アイテムについての予測評価値を求めると、予測評価値算出部104は、未評価アイテムがある他のユーザを特定のユーザとし、上述と同様な処理 (S1~S4) を繰り返して、全てのユーザの未評価アイテムについて予測評価値を求める。

【0063】全てのユーザの未評価アイテムの予測評価値が求められると、アイテム推薦部106は、ユーザaに、例えば予測評価値が高い順に所定数のアイテムを推薦する。あるいは、予め所定の閾値を決めておき、予測評価値が当該閾値を超えたアイテムを推薦するようにしてもよい。

【0064】具体的なアイテムの推薦は、上述のように、例えばSMTPサーバ107による当該ユーザ宛て

11

のメールの送信あるいはHTTPサーバ103による推薦するアイテムをすべ記述データの提供等によって行う。

【0065】アイテム推薦部106は、メールによってアイテムを推薦する場合に、例えば推薦するアイテムを示す文字、当該アイテムのアドレスの記述等を含むメールのメッセージを生成し、SMTPサーバ107によってユーザ宛てに送信させる。また、ページ記述データによってアイテムを推薦する場合に、アイテム推薦部106は、例えば推薦するアイテムを示す文字、当該アイテムのアドレスの記述等を含むページ記述データを生成し、当該ユーザのユーザIDと共に、HTTPサーバ103に供給する。HTTPサーバ103は、例えば当該ユーザからのアクセスがあると、アイテム推薦部106から供給されたページ記述データを当該ユーザの携帯端末装置3宛てに送信する。

【0066】（効果）上述のように、この推薦システムでは、ユーザによるアイテムのアクセス履歴、ダウンロード履歴等に基づいて求めた暗黙的な評価データを用いて、ユーザに推薦するアイテムを決定しているため、ユーザの負担は非常に軽い。

【0067】また、この推薦システムでは、ユーザ等に関与するアイテムに対するアクセス時刻からの時間の経過に従って減少する評価値を求めることにより、アクセス履歴、ダウンロード履歴に相関の高い評価値を求めることができる。

【0068】このような評価値を用いてユーザに推薦するアイテムを決定することにより、ユーザの傾向をより正確に反映したアイテムの推薦を行うことができ、推薦フィルタリングの精度を向上させることができる。

【0069】また、上述のように、この推薦システムでは、ユーザにアイテムを推薦する際に、推薦するアイテムにアクセスするための情報（例えば推薦するアイテムのアドレス等）をユーザに提供することにより、特に、携帯電話装置、PDA等の表示能力が制限された携帯端末装置を用いて、コンテンツをダウンロードする際の手数を大幅に削減することができる。

【0070】（変形例）上述の説明では、上述の図5に示すように、評価データ作成部104が予めアクセスログ101bに記憶したユーザ等の個々のアイテムの評価値を求めておくようにしていたが、例えば図7に示すように、評価データ作成部104は、単に、アクセスログ101bからユーザ等の個々のアイテムに対するアクセスの有無及びアクセス時刻を評価データベース101cに格納しておくようにしてもよい。なお、この図7は、説明の便宜上、上述の図5に示す状況とは完全には対応していない。

【0071】この場合は、予測評価値算出部105が予測評価値を求める際に、現在の時刻と評価データベース101c中のアクセス時刻に応じて上述の数1式により

(7)

特開2003-167901

12

係数を求め、求めた係数を乗じた評価値に応じて、予測評価値を求める。

【0072】例えばこの図7に示す状況では、ユーザ1の評価値の傾向とユーザn-2等の評価値の傾向が似ている。このため、ユーザn-2等が、ユーザ1と相関が高いユーザとして選択される。この後、ユーザ1が評価していないアイテムnに對する予測評価値が、上述と同様に、ユーザn-2等がアイテムnにアクセスした時刻に応じた値として求められる。

【0073】図7に示す状況では、ユーザn-2等がアイテムnにアクセスした時刻が比較的最近である。このため、ユーザn-2等がアイテムnにアクセスした時刻に応じて上述の数1式から求めた係数が大きくなっていく。これにより、ユーザ1の予測評価値が大きくなるため、ユーザ1が近い将来、アイテムnにアクセスすると予測することができる。

【0074】なお、上述の説明では、アイテムの例として書信メロディを示して説明したが、本発明はこれに限定されるものではなく、例えばニュース等の文字情報、着信メロディ、待ち受け画面、音楽等のコンテンツ（アイテム）を推薦する推薦システムに本発明を適用することができる。これらのアイテムを提供する際には、例えば上述のHTTPサーバ103あるいは特別なサーバプロセスを用いるが、上述と同様なアクセスログを記録しておくことにすれば、このアクセスログに応じて上述と同様にアイテムの推薦を行うことができる。

【0075】また、上述の説明では、無線通信を介してサーバ装置に接続された携帯端末装置によってサーバ装置が提供するアイテムを利用する構成とした例について示した。しかしながら、本発明は、この例に限定されるものではなく、例えば有線ネットワークを介してサーバ装置に接続されたパーソナルコンピュータ、ワークステーション等の情報処理装置によってサーバ装置が提供するアイテムを利用する構成とする等、本発明の技術的思想の範囲内で、適宜変換を要することがある。

【0076】

【発明の効果】本発明では、ユーザの行動履歴に応じた評価対象についての評価を示す暗示的评价データと、ユーザの行動による評価の時点からの経過時間に応じた評価係数を求め、特定のユーザと評価情報の傾向が類似するユーザを選択し、選択したユーザが評価している評価対象についての評価情報に応じて、特定のユーザの当該評価対象についての予測評価値を求めるため、従来の推薦フィルタリングにおける評価データでは考慮されていなかった時間情報を考慮することができる。

【0077】同じような行動をしても当該行動の時期が異なるユーザは、異なる傾向を有すると考えられるため、上述のように時間情報を考慮して求めた評価情報を用いることにより、より傾向の近いユーザを選択することができる。

(8)

特開2003-167901

13

【0078】のように選択したユーザの評価情報に応じて、特定のユーザの予測評価値を求めることにより、求められる予測評価値の精度を向上させることができる。

【0079】また、本発明では、ユーザの行動履歴に応じた評価対象についての評価を示す暗示的評価データに基いて予測評価値を求めているため、ユーザは意識的に評価を行う必要がない。このため、ユーザの負担を軽減することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る総務システムの構成を示すブロック図である。

【図2】前記総務システムを構成するHTTPサーバが生成するアクセスログの一例を示す図である。

【図3】前記総務システムを構成する評価データ作成部が用いる係数の一例を示す図である。

【図4】前記評価データ作成部が作成する評価データベースの一例を示す図である。

19

*【図5】前記評価データ作成部が作成する評価データベースの一例を概念的に示す図である。

【図6】前記評価システムを構成するアイテム総務部による予測評価値を求める処理を示すフローチャートである。

【図7】前記評価データ作成部が作成する評価データベースの他の例を概念的に示す図路図である。

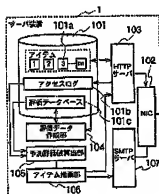
【図8】従来の協調フィルタリングにおいて用いられる明示的な評価データの例を示す図である。

【図9】従来の協調フィルタリングにおいて用いられる暗示的な評価データの例を示す図である。

【符号の説明】

1…サーバ装置、101…HDD、101a…アイテム格納領域、101b…アクセスログ、101c…評価データベース、102…NIC、103…HTTPサーバ、104…評価データ作成部、105…予測評価値算出部、106…アイテム総務部、107…SMTPサーバ

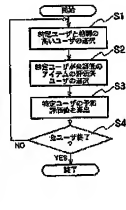
【図1】



【図2】

101b				
	ID	アクセス時刻	アクセス元	ユーザID
a	0001	01/10/30	XX.XX	1
	0002	01/10/31	YY.YY	2
b	0003	01/10/31	YY.YY	2
	0004	01/11/05	XX.XX	1
c	0100	01/11/05	XX.XX	1
	0101	01/11/06	YY.YY	2

【図6】



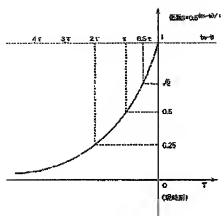
【図4】

101c				
アイテム		ユーザ別		
	ユーザID=1	2	...	n-2
1	0.3	0.4	0.5	0.6
2	0.4	0.5	0.6	0.7
3	0.5	0.6	0.7	0.8
4	0.6	0.7	0.8	0.9
5	0.7	0.8	0.9	1.0
...
m

(9)

特開2003-167901

【図3】



【図5】

アイテム						
	1	2	3	4	5	m
(ユーザ1)	0.5	0.4	0.8	0.5	0.2	0
(ユーザ2)	0.4	0.5	0.7	0	0.2	1.0
(ユーザ3)	0	0	0	0.5	0	0
(ユーザ4)	0.5	0.2	0.1	0	0.7	0.2

【図7】

アイテム						
	1	2	3	4	5	m
(ユーザ1)	2001.6.1	2001.6.5	0	2001.6.7	2001.6.10	0
(ユーザ2)	2001.6.20	2001.6.22	2001.6.23	0	2001.6.22	2001.6.1
(ユーザ3)	0	0	0	2001.6.20	0	0
(ユーザ4)	2001.6.22	2001.6.25	0	2001.6.1	2001.6.1	2001.6.1

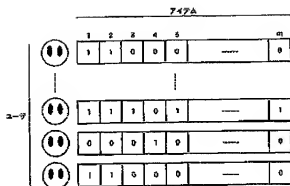
【図8】

アイテム						
	第1項	第2項	第3項	第4項	第5項	第6項
(ユーザ1)	5	1	-	2	3	-
(ユーザ2)	2	-	1	3	3	2
(ユーザ3)	-	2	4	1	-	5
(ユーザ4)	3	4	3	-	5	-

(10)

特開2003-167901

【図9】



フロントページの続き

(72)発明者 田中 卓弥
 埼玉県上福岡市大原2丁目1番15号 株式
 会社ケイディーディーアイ研究所内

(72)発明者 井ノ上 直己
 埼玉県上福岡市大原2丁目1番15号 株式
 会社ケイディーディーアイ研究所内

Fターム(参考) 5B075 NK06 PR03 PR08